2

Information Item

California Postsecondary Education Commission

Executive Director's Report, December 2000

Executive Director Warren Fox will discuss issues of mutual concern to the commissioners. Included in this report will be Commission information about the report, *Measuring Up 2000: The State-By-State Report Card for Higher Education*, published by the National Center for Public Policy and Higher Education.

In addition, Director Fox will discuss issues related to the numerous testing efforts underway in California to assess student learning in both K-12 and higher education.

Presenter: Warren H. Fox, Executive Director.

CALIFORNIA POSTSECONDARY EDUCATION COMMISSION

Report of the Executive Director

December 11, 2000

Introduction

In recent years, the Commission has called for greater accountability as it relates to the educational services provided to students -- an effort intended to improve the quality of education that students ultimately receive. The first item in this report discusses a newly issued report card for state-by-state performance in higher education, which illustrates the responsibility states have in ensuring that their citizens are provided with quality educational opportunities. The second topic is the critical issue of student testing in California secondary and postsecondary educational systems.

"Report card" on higher education in the states

As discussed at the August Commission meeting, the National Center for Public Policy and Higher Education in San Jose is issuing a state-by-state "report card" on higher education. It is expected that *Measuring Up 2000: The State–By-State Report Card For Higher Education* will be released on December 1, 2000.

The report compares state-by-state performance in higher education. It is anticipated that letter grades will be assigned for each state in six issue areas: (1) Preparation, (2) Participation, (3) Affordability, (4) Completion, (5) Benefits, and (6) Learning. The National Center indicates that its intention in the initial report card is to emphasize the "traditional guiding values of state policy for education" as well as to grade states in comparison with what it considers to be the highest performing states. The authors have stated that their hope is to stimulate the creation of state policies that enhance opportunity and achievement for all Americans who aspire to higher education. The report card is intended to serve as a diagnostic tool for state policy leaders to enable them to identify the strengths and weaknesses of higher education performance in their state compared to others and to serve as catalyst for constructive change as a framework for discussion and debate.

Once available, staff will share information gleaned from the "report card" with the Commission. There has been anxiety among some higher education leaders across the nation about the use these report cards and others like them. While it can agreed that these are less-than-precise indicators of the effectiveness of a state's higher education system, such indicators can be useful in helping provoke discussion and illuminating both strengths and weaknesses in higher education systems. It is Commission staff's perspective that this particular report card could be a useful tool to help us either reinforce or initiate creative strategies as plans are made for the future. Among the questions to consider:

- Do we agree with the report card's assessment of our effectiveness in each of the areas, and if not, why not?
- Are we doing an adequate job of addressing each of these major areas, as well as others we believe to be important?
- What can the Commission, as California's higher education planning and coordinating board, do to better address these issues?
- What can we learn from other states who are perceived to be more adequately addressing a particular need than California? Are there models that can be adapted here in California?

It is anticipated that the "report card" will be provided to members of the Commission at the December meeting for discussion and comment.

Testing issues in secondary and postsecondary education in California

In recent years, testing of students has increasingly become a high-stakes enterprise in California. Concerns about low student achievement scores have prompted calls for tougher standards and evaluation methods that attempt to quantify quality that enable policy makers to hold educators and students more accountable for meeting those standards.

In California, as in many other states, testing offers both a "carrot-and-stick" approach to accountability. Teachers and schools receive incentive funding for rising test scores, and are labeled "low performing" for test scores that are below average. These high stakes extend to the world beyond school as well.

As California enters the 21st century, a technology-driven economy will require a workforce with advanced skills creating very real "high stakes" for California's economy. Many believe that, in order to remain competitive in the new economic environment, California must ensure that future workers will possess the knowledge and skills they will need be productive citizens and successful individuals. As more has come to be expected of students, teachers, and schools, policy makers are looking to various forms of testing as a way to measure how well these groups are meeting those higher, tougher expectations.

Yet the issue of testing is not the sole domain of those involved in California's elementary and secondary schools. Not only is the higher education enterprise in California dealing with similar issues of measuring student achievement and providing greater accountability as K-12, there is also the critical issue of ensuring the establishment of a more natural and seamless connection between the testing efforts of K-12 and higher education.

Testing students for various purposes from kindergarten through university has enormous implications for eligibility, remediation, educational equity, and educational opportunity for all students. The Commission has long had an interest and involvement in these issues. This document is intended to provide the Commission with some background information on the history and current uses of testing in California as well as to provide a brief summary

of some of the issues at stake in the debate over assessment. It is intended to inform the Commission's future policy work in this area.

History of testing

Testing has long been a part of schools and education. During the Middle Ages, oral examinations were common for some disciplines in European universities. Oral examinations for the Bachelor's and Master's degree were first used at Oxford in 1636 (Popham, 1990). In early America's one-room schoolhouses, tests were primarily oral and were probably used to determine if students were ready to advance to more challenging studies. In their most basic form, tests provide a way for students to demonstrate what they know and for teachers to gauge how much students have learned.

The first known effort for American schools to test students in a consistent, standardized manner occurred in 1845 when the Boston school system began using printed short-answer tests. Such ventures provided educators the opportunity to understand the complexity of the teaching and learning relationship and the demonstration of knowledge.

Pioneer efforts of more modern testing began at the turn of the century, with Joseph Mayer Rice and Edward Thorndike among the first to use standardized tests to establish average scores for different grade levels. Thorndike is believed to be the first to establish achievement norms for a variety of subjects, including arithmetic, reading, and handwriting. By administering a test to a large number of students and summarizing the average performance of the "norm group," Thorndike's norm-referenced methodology allowed educators, for the first time, to compare their students' performance with the average performance of other students. As others emerged, so did new forms of standardized tests that measured innate ability and allowed educators to "predict" future performance.

In 1905, French psychologist Alfred Binet created a scale he claimed could ascertain whether a child would benefit from instruction. By being able to estimate a child's "mental age," Binet's test claimed it could identify a student's capacity to learn. A Stanford University psychologist, Lewis M. Terman, took Binet's work a step further by dividing the "mental age" by the chronological age to arrive at what Terman called the Intelligence Quotient, or "IQ."

The resulting Stanford-Binet scale, published in 1916, provided instruction on the use of the Binet test to assess a student's inherent intelligence and capacity to learn. The first wide-scale use of intelligence tests to determine individual aptitude came as the United States prepared for World War I. The U.S. Army, looking for a way to identify potential officer candidates and make more effective staffing decisions, administered multiple-choice intelligence tests to more than 1.7 million men. The success of the Army's testing program led to a proliferation of standardized multiple-choice item aptitude tests and, by 1922, these IQ tests were used frequently by educators to classify and track students in school.

Dr. Terman went on to develop the Stanford Achievement Test for the purpose of measuring student performance. It was normed using a large national sample of 350,000 students and became the model for standardized tests.

Although entrance exams had long been used by some colleges, Harvard University President James Conant's interest in establishing an objective method of evaluating applicants and selecting scholarship students led to the use of the Scholastic Aptitude Test (SAT) in college admission decisions. Now called the Scholastic Achievement Test, the SAT is an integral part of the college application process and a rite of passage for the more than two million students who take the exam each year.

In addition to their utility in assessing knowledge and abilities of individuals, test data can provide educators and policy makers with information about the aggregate performance of students from which conclusions can be drawn about the overall quality of teaching and educational services in a given school setting. Although there are a number of ways to evaluate student achievement, standardized tests are often chosen because of their reliability and means of providing an objective assessment of a student's performance. Their consistent administration and scoring procedures enable evaluators to interpret scores in a uniform manner and compare results of different test takers.

As California and the nation move into the 21st century, the demand for greater school quality has led to more reliance on test data as a gauge of progress and accountability. The publication of "A Nation at Risk" in 1983 has been cited by some as a catalyst in the call for greater accountability in public schools. The low student-achievement scores reported in it alarmed parents, educators, and policy makers, resulting in a call for tougher academic standards and stronger, better schools. The measurement of school quality shifted from the tabulation of resources to the assessment of student learning. Policy makers began to look at standardized tests as a means of measuring student progress and evaluating educational quality.

Types of testing

There are two major categories of tests: Norm-referenced tests and criterion-referenced tests. While both may be "standardized" the main differences between the two types of tests can be found in their intended purpose and in the way test scores are interpreted.

Norm-referenced tests are the most common and are used to compare student performance with a "norm" or average performance of a sample peer group. The content and criteria may or may not reflect any particular curriculum or academic goals established by governing boards or policy makers. Because norm-referenced tests are typically "normed" against large national or state samples of test takers, their content usually reflects generalized knowledge or skill sets. Examples of norm-referenced tests include the SAT I and the Stanford 9 test used in California public schools.

Scores from norm-referenced tests are interpreted by indexing individual performance relative to the average performance of students in a comparison

group, and are often expressed in terms of standing in a percentile group. For example in 1998, a score of 1220 on the SAT I was in the 75th percentile, meaning that the student's performance was better than 75 percent of the students in the norm group.

The utility of such data lends itself to sorting or rank ordering students and often provides an objective basis for determining and addressing a student's educational needs. Top scoring students might be selected for enrichment programs, like California's Gifted and Talented Education (GATE) program, while those scoring below 50 percent might be targeted for extra help in specified subject areas. Scores from a norm-referenced test provide information about how a student performs relative to other students but do not provide information about how much an individual student has learned relative to specified achievement standards.

Criterion-referenced tests, in contrast, do provide information about how individuals perform relative to defined objectives. These tests help educators ascertain what test takers can do and what they know. Because these tests provide information about how well students perform relative to defined learning outcomes instead of a comparison group, they are useful in determining the progress students are making toward mastering the knowledge and skill levels expected of them. The content and criteria of a criterion-referenced test reflect the curriculum and items are selected on how well they represent what students should know and be able to do. The greater the congruence between the test content and the curriculum, the more the test will provide valid information about the performance of students and teachers. Criterion-referenced tests are based on standards.

In the early 1990s, the assessment movement attempted to bridge the gap between these basic test styles by attempting to broaden tests to include open-ended questions and developing less structured and less standardized approaches to gauging student achievement. Critics argue that multiple-choice questions, the standard format for many tests, do not assess a student's ability to come up with his or her own answers.

Performance-based assessments, which help educators look at how test takers demonstrate competency and portfolios, which provide an assessment of growth over time, are two popular alternative assessment methods. These assessments tend to provide a more "complete" picture of student performance, allowing evaluators to understand responses in context. However, these assessments are often expensive to administer and are often fraught with questions about measurement error and bias. The California Learning Assessment System (CLAS), for example, met with public outcry about test questions and reading material and was ultimately abandoned as a public school testing program.

In order for tests to be effective measurement devices, they must be used for their intended purpose and they must provide reliable and valid measurements. If the question is "How do our students compare with other students in basic skill areas?" a norm-referenced test is appropriate. If, on the other

hand, the question is, "How well are our students learning what we want them to learn?" a criterion-referenced test would be the better choice. Whether they are norm-referenced or criterion-referenced, good tests must be both consistent in their ability to measure performance and accurately represent what they claim to measure.

K-12 testing in California

California public schools tests attempt to answer questions about how students compare to others, what they know and are able to do, and how well schools carry out instruction. New K-12 content and performance standards and the public demands more accountability for the investment of tax dollars in education helps ensure that testing assumes a prominent role. There are now more than 10 different exams that test achievement, proficiency, college eligibility, and determine placement. These includes:

Achievement Tests
Stanford 9 (STAR Program)
Golden State Examination
High School Exit Exam

College Entrance Exams
PSAT
ACT
SAT I
SAT II Subject Tests

Proficiency/Placement Tests
Entry Level Mathematics Exam
English Placement Test
TOEFL
Subject A
AP/IB Exams
MDTP

The Stanford 9/ STAR program

Adopted in October 1997 after looking at various approaches to assessment, the Standardized Testing and Assessment Reporting (STAR) program. Enacted by SB 376, it required the California State Board of Education to develop academic content standards and the systematic testing of nearly all public school children. The State Board designated the *Stanford Achievement Test* series for the STAR program. The norm-referenced SAT 9 assesses a broad range of basic academic skills and provides comparable individual pupil scores. Students in grades 2 through 8 are tested in reading, mathematics, writing, and spelling. Students in grades 9 through 11 are tested in reading, writing, mathematics, science, and history/social science. The first administration of the exam occurred in the spring of 1998. In 1999, the SAT 9 was augmented to include questions aligned with recently adopted academic content standards.

Reported statewide and by county, district, and school, the results help to answer the question about "how well schools carry out instruction." The Public School Performance Accountability Act, established by SB 1X (Alpert, Statutes of 1999), created rewards and interventions for schools and educators as a means of improving student performance. The act also provides an "Immediate Intervention" grant program for helping underperforming schools and an incentive grant program for high achieving/improving schools.

Golden State Exam Program

The Golden State Examination (GSE) program is a voluntary examination program available to students in grades 7-12 providing college-level examinations in 18 subject areas and provides a means of recognizing students who demonstrate outstanding levels of achievement in several academic subject areas. Districts are required to make Golden State Examinations available to all students. Students who attain the three highest levels of achievement designations (high honors, honors, or recognition) on six Golden State Examinations are eligible to receive a Golden State Seal Merit Diploma upon their graduation from high school (less than 1% of California's 1998 high school graduates did so).

High School Exit Exam

California is in the process of establishing higher standards for high school graduation and developing a High School Exit Examination (HSEE). The purpose of the exam is to "ensure that pupils who graduate from high school can demonstrate grade level competency in reading, writing, and mathematics." Senate Bill 2X (1999) requires that, beginning in the 2003-2004 school year, students will not receive a high school diploma without first achieving a passing score on an exit examination. Beginning in 2000-01, 9th grade students will be eligible to take the exam. In 2001-02, 10th grade students will be required to take the exam. These students may take the exam during each administration until they pass each section. The exam is to be consistent with the State content standards for language arts and mathematics.

Higher education assessments

Students not only are subject to the three achievement tests listed above, but there are also myriad tests for students pursuing higher education. In general, they serve as either a tool for admissions purposes or for placement. Below is a brief description of some of the major tests used by higher education institutions in California.

Scholastic Assessment Test I (SAT I): This is a norm-referenced, multiple-choice exam that measures mathematical and verbal reasoning. It is used for college selection and is intended to predict success in college. Both UC and CSU utilize the SAT I. All applicants to UC are required to submit SAT I scores, although a student may choose to take the ACT. Applicants to CSU with less than a 3.0 high school grade point average must submit their SAT or ACT scores for admission.

Scholastic Assessment Test II (SAT II): This is a set of one-hour, norm-referenced, multiple choice and open ended tests used primarily for admission to the more selective colleges and universities. Proponents of these tests suggest they are more closely aligned to the academic content in particular subject areas and therefore better predictors of student achievement. The University of California requires that to be eligible for admission, all applicants must take three SAT II tests. In addition, the University has recently begun assigning greater weight to the SAT II tests than it had previously in determining admission.

The American College Testing (ACT): This exam is a three-hour, multiple-choice exam that assesses achievement in several academic subjects. It is

used primarily used for college admission. Although the vast majority of students in California take the SAT rather than the ACT, approximately 38,000 California students took the ACT in 2000. Both CSU and UC accept the ACT as well as the SAT.

Test of English as a Foreign Language (TOEFL): This test is for students who have not attended schools at the secondary level or above for at less three years full time where English is the principal language of instruction. It is intended to measure English proficiency.

Advanced Placement (AP) Exams: These exams are used to measure college-level achievement in many different subject areas and to award academic credit to high school students who demonstrate college-level proficiency. Scores ranging from 1 (No recommendation) to 5 (Extremely well qualified) are awarded to students who take the exams. These scores are used as evidence of the students' abilities and achievements and allow colleges and universities to make their decisions regarding whether or not to grant credit and or/advanced placement.

Entry Level Math (ELM) Exam: This examination is designed to assess the skill levels of entering CSU students in the area of mathematics typically covered in the three years of rigorous college preparatory mathematics courses in high school. Those undergraduate students who do not demonstrate college-level skills will be directed to courses or programs designed to help them attain these skills. Most entering undergraduates at CSU take the ELM examination before enrolling in a course that satisfies the college-level mathematics requirement of the General Education-Breadth program. The exceptions are those who score sufficiently high on the mathematics section of the SAT I: Reasoning Test, SAT II: Mathematics Test, ACT Mathematics Test, Advanced Placement Mathematics Examination (Calculus AB or BC), Advanced Placement Statistics Examination, or transfer students who have satisfied the Quantitative Reasoning requirement under the Intersegmental General Education Transfer Curriculum.

English Placement Test (EPT): This test, used by the CSU system, is designed to assess the level of reading and writing skills of entering undergraduate students so that they can be placed in appropriate baccalaureate courses. The CSU English Placement Test must be completed by all entering undergraduates with the exception of those who score sufficiently high on the SAT Reasoning Test, the SAT II Writing Test, AP Language and Composition or Literature Exam, or transfer students who have completed the English Composition requirement under the Intersegmental General Education Transfer Curriculum.

Subject A Exam: The University of California requires entering students who do not meet the University of California standards for English to take a two-hour essay exam. Students are required to read a passage and then write an essay responding to a single topic based on the content of the passage. All students who enter the University of California directly from a California high school must take the University-wide Subject A Examination unless

they have a score of 680 on the SAT II Writing Test or a 3 or higher on the Advanced Placement Examination in English. Students may take the exam only once and those who do not pass can fulfill the requirement by either achieving a satisfactory SAT II Writing or AP English score, complete with a C or better an acceptable college course in English composition, or achieve a score of 5 or better on the International Baccalaureate's Higher Level English A Examination. Students who have not satisfied the Subject A requirement prior to enrolling in the University must take and pass with a C or better, a writing course designated by their campus for satisfying the Subject A requirement.

Mathematics Diagnostic Testing Program (MDTP): The University of California has no official system-wide policy on mathematics placement; however, many UC campuses use the Mathematics Diagnostic Testing Program (MDTP). The program was developed jointly by CSU and UC faculty members and consists of five tests: Algebra Readiness, Geometry Readiness, Second Year Algebra readiness, Mathematical Analysis Readiness, and Calculus Readiness. These tests are administered by high school and middle school teachers.

Diagnostic Writing Services (DWS): This service is available for students and teachers to assess current writing skills against college-level expectations. It allows individual students to write an essay in response to an actual English Placement Test essay prompt and to submit the essay via the Internet. A university EPT reader uses the scoring rubric of the EPT test and assesses student writing against the standards expected of entering college students. Diagnostic statements are provided to the student via the Internet.

Writing Proficiency Exam: All candidates for a bachelor's degree at CSU must take and pass, prior to graduation, a test that assesses a student's ability to write with college-level proficiency. The student must compose an essay from a topic that is provided to them.

Current efforts

The California Education Roundtable has long recognized the importance of collaboration and cooperation between K-12 and postsecondary education to ensure that assessment of students throughout the educational continuum is better aligned and streamlined.

To that end, the Education Roundtable has convened a working group to begin to develop policy principles on the issue of alignment and to formulate a more logical and articulated testing and assessment system for California's students. The Commission will be participating in these meetings which will begin in mid December and staff will provide the Commission with updates on the progress of this group.

Conclusion

As policy makers continue to consider ways to improve our educational system, strengthen the quality of instruction provided to our students, and raise achievement levels, there will likely be many initiatives aimed at measuring the extent to which we are in fact meeting the new standards and higher expectations that have been established for students. Used appropriately, test-

ing can be a tool for demonstrating improvement, measuring the quality of instruction, and determining gains in student achievement across the educational continuum.

California has invested vast resources in recent years to creating uniformity in the content of our curriculum, reducing class sizes, strengthening teacher quality, and setting higher standards for educators and their students. It is important that the process selected to evaluate these efforts informs the State's ability to continue the course toward better educational institutions – elementary through university -- stronger instruction, and successful students.

The Commission will continue to monitor the activities underway in California on the issue of testing, participate in discussions intended to enhance the alignment of K-12 and higher education assessments, explore the policy implications associated with the topic, and provide recommendations for Commission consideration, where appropriate.